Histone and TK0471/TrmBL2 form a novel heterogeneous genome architecture in the hyperthermophilic archaeon Thermococcus kodakarensis.

AUTHOR(S):

Maruyama, Hugo; Shin, Minsang; Oda, Toshiyuki; Matsumi, Rie; Ohniwa, Ryosuke L; Itoh, Takehiko; Shirahige, Katsuhiko; ... Atomi, Haruyuki; Yoshimura, Shige H; Takeyasu, Kunio

CITATION:

Maruyama, Hugo ...[et al]. Histone and TK0471/TrmBL2 form a novel heterogeneous genome architecture in the hyperthermophilic archaeon Thermococcus kodakarensis.. Molecular biology of the cell 2011, 22(3): 386-398

ISSUE DATE:

2011-02

URL:

http://hdl.handle.net/2433/138088

Histone and TK0471/TrmBL2 form a novel heterogeneous genome architecture in the hyperthermophilic archaeon *Thermococcus kodakarensis*

Hugo Maruyama[a], Minsang Shin[a], Toshiyuki Oda[a], Rie Matsumi[b], Ryosuke L. Ohniwa[c], Takehiko Itoh[d], Katsuhiko Shirahige[e], Tadayuki Imanaka[f], Haruyuki Atomi[g], Shige H. Yoshimura[a], and Kunio Takeyasu[a]

[a]Laboratory of Plasma Membrane and Nuclear Signaling, Graduate School of Biostudies, Kyoto University, Kyoto 606-8501, Japan; [b]Laboratory of Microbiology, Wageningen University, 6703 HB Wageningen, The Netherlands; [c]Institute of Basic Medical Sciences, Graduate School of Comprehensive Human Sciences, University of Tsukuba, Tsukuba 305-8575, Japan; [d]Laboratory of In Silico Functional Genomics, Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, Yokohama 226-8501, Japan; [e]Laboratory of Genome Structure and Function, Research Center for Epigenetic Disease, Institute of Molecular and Cellular Biosciences, The University of Tokyo, Tokyo 113-0032, Japan; [f]Department of Biotechnology, College of Life Sciences, Ritsumeikan University, Kusatsu 525-8577, Japan; [g]Department of Synthetic Chemistry and Biological Chemistry, Graduate School of Engineering, Kyoto University, Kyoto 615-8510, Japan

ABSTRACT **Being distinct from bacteria and eukaryotes, Archaea constitute a third domain of living things. The DNA replication, transcription, and translation machineries of Archaea are more similar to those of eukaryotes, whereas the genes involved in metabolic processes show more similarity to their bacterial counterparts. We report here that TK0471/TrmB-like 2 (TrmBL2), in addition to histone, is a novel type of abundant chromosomal protein in the model euryarchaeon *Thermococcus kodakarensis*. The chromosome of *T. kodakarensis* can be separated into regions enriched either with histone, in which the genetic material takes on a "beads-on-a-string" appearance, or with TK0471/TrmBL2, in which it assumes a thick fibrous structure. TK0471/TrmBL2 binds to both coding and intergenic regions and represses transcription when bound to the promoter region. These results show that the archaeal chromosome is organized into heterogeneous structures and that TK0471/TrmBL2 acts as a general chromosomal protein as well as a global transcriptional repressor.**

Monitoring Editor
William P. Tansey
Vanderbilt University

Received: Aug 4, 2010
Revised: Nov 12, 2010
Accepted: Nov 23, 2010

This article was published online ahead of print in MBoC in Press (http://www.molbiolcell.org/cgi/doi/10.1091/mbc.E10-08-0668) on December 9, 2010.

Author contributions: H. Maruyama, M. Shin, and R. Matsumi designed and performed experiments. K. Shirahige and T. Itoh performed and analyzed massively parallel sequencing of DNA. T. Oda and R. L. Ohniwa performed bioinformatics analyses. T. Imanaka, H. Atomi, and K. Takeyasu conceived and supervised the project. H. Maruyama, R. L. Ohniwa, S. H. Yoshimura, and K. Takeyasu interpreted the data and wrote the paper.

The authors declare they have no conflict of interest.

Address correspondence to: Hugo Maruyama (maruyama@lif.kyoto-u.ac.jp).

Abbreviations used: AFM, atomic force microscopy; Alba, acetylation lowers binding affinity; ASW, artificial seawater; DAPI, 4′,6-diamino-2-phenylindole; emPAI, exponentially modified protein abundance index; HSD, honestly significant difference; HTH, helix-turn-helix; MNase, micrococcal nuclease; NAP, nucleoid-associated protein; ORF, open reading frame; RNAP, RNA polymerase; RPA, replication protein A; SG, sucrose gradient; TrmB, transcriptional regulator of mal operon.

INTRODUCTION

In eukaryotic cells, genomic DNA is constrained into nucleosomes by histone proteins (Luger *et al.*, 1997). The array of nucleosomes is further organized into higher-order structures, which play important roles in processes such as transcription, replication, and chromosome segregation (Ehrenhofer-Murray, 2004; Maeshima and Eltsov, 2008). The bacterial nucleoid is likewise folded into stepwise higher-order structures (Kim *et al.*, 2004; Ohniwa *et al.*, 2006; Thanbichler and Shapiro, 2006), although its protein components are fundamentally different from those of eukaryotes (Dillon and Dorman, 2010; Ohniwa *et al.*, 2010). Among the bacterial nucleoid-associated proteins (NAPs), HU is the most highly conserved; it is considered to play critical roles in the construction of nucleoids. Other proteins (e.g., IHF, Fis, and H-NS) are also involved in structuring the *Escherichia coli* nucleoid (Ali Azam *et al.*, 1999), but these are not conserved in other bacterial phyla such as Firmicutes (Kim *et al.*, 2004).

Archaea are distinct from both bacteria and eukaryotes and thus constitute a third domain of living things. This domain is currently divided into two main phyla, Crenarchaeota and Euryarchaeota, along with several other newly proposed phyla (Woese *et al.*, 1990; Gribaldo and Brochier, 2009). The archaeal species possess a mixture of bacterial and eukaryotic features. For example, the DNA replication, transcription, and translation machineries of Archaea are more similar to those of eukaryotic organisms, whereas the genes involved in metabolic processes show more similarity to their bacterial counterparts (Koonin *et al.*, 1997; Smith *et al.*, 1997).

Archaea are interesting organisms to use in the study of chromosome organization because of the diversity of chromosomal proteins among species. Most crenarchaeal species encode a protein called Alba (acetylation lowers binding affinity; also known as Ssh10b, Sso10b, and Sac10b, depending on species of origin), whose affinity to DNA increases upon deacetylation by the silencing protein Sir2 (Bell *et al.*, 2002). Alba can constitute up to ~4% of the total cellular proteins in *Sulfolobus* (Xue *et al.*, 2000) and is considered a major chromosomal protein. Some euryarchaeal species also encode Alba, although the protein is less abundant in these cells. Most euryarchaeal species, interestingly, encode a protein homologous to eukaryotic histone, with some exceptional species that encode a protein homologous to bacterial HU. The archaeal histones consist solely of the histone-fold motif (Sandman and Reeve, 2005) and lack the N- and C-terminal tails that are posttranslationally modified and involved in chromatin dynamics in eukaryotes (Kouzarides, 2007). Unlike eukaryotic histones, which form octamers, archaeal histones form tetramers (Pereira *et al.*, 1997) around which ~120 base pairs of DNA can wrap in vitro (Tomschik *et al.*, 2001). Although many studies have supported the conclusion that archaeal histones form nucleosome structures in vivo (Shioda *et al.*, 1989; Pereira *et al.*, 1997), whether and how other proteins contribute to the chromosome architecture in these species is largely unknown.

An understanding of the organization and dynamics of the archaeal chromosome may provide new insight into the presence or absence of a common folding mechanism among all three domains of life. In the present study, we identified the protein components of the chromosome of *Thermococcus kodakarensis*, a histone-coding euryarchaeon, and elucidated how its overall chromosome structure is organized through the coordination of different proteins.

RESULTS

T. kodakarensis chromosome consists of beads-on-a-string structures and fibrous structures

Hypotonic treatment of *T. kodakarensis* KOD1 cells attached to a cover glass can disrupt the cell membrane and release the chromosome fibers from the cells. The chromosome fibers exposed from the cells in the log and stationary phases were visualized by staining DNA with 4′,6-diamino-2-phenylindole (DAPI). The degree of chromosome spreading was greater in the cells from the log phase culture than in those from the stationary phase (Figure 1A). Detailed results were obtained through structural analysis by atomic force microscopy (AFM) (Figure 1B). Chromosome fibers released from the cells in the log phase exhibited both a beads-on-a-string structure and fibrous structures in the regions where the chromosome fibers were well spread. The diameter of the "bead" was 8.7 ± 1.1 nm (mean ± SD, n = 150) (Figure 1C). This size was similar to that of archaeal

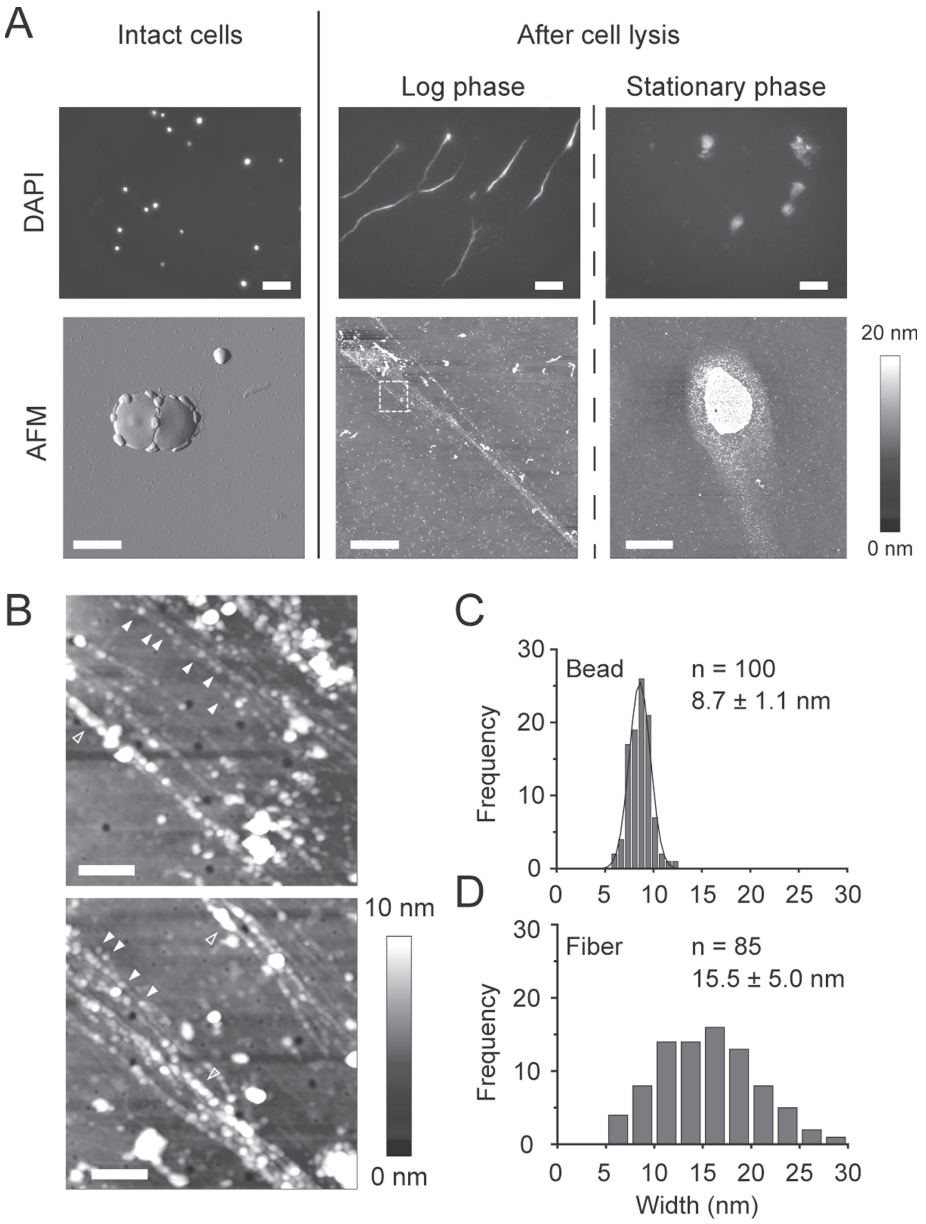


FIGURE 1: Chromosome fiber of *T. kodakarensis* analyzed through AFM. (A) DAPI staining (top; scale bars, 10 μm) and AFM images (bottom; scale bars, 2 μm) of chromosomes released from *T. kodakarensis* cells in the log and stationary phases. Intact cells are shown on the left. (B) Detailed AFM analysis of the chromosome fiber. Beads-on-a-string structures (solid triangles) as well as thick structures (open triangles) were observed in areas where chromosomes were well spread out (boxed area in [A]). Scale bars, 200 nm. Histograms indicate the diameters of (C) beads and (D) fibrous structures. The means ± SD of the distributions are indicated. The curve indicates the fitted Gaussian function.

nucleosomes isolated from the chromosome or reconstituted with recombinant histone protein (see below). The diameter of the fibers had a broader distribution (15.5 ± 5.0 nm, n = 85) (Figure 1D). Although some portion of the thick fibers could have resulted from bundling of chromosome fibers released upon cell lysis, these results indicate the existence of heterogeneous structural elements in the *T. kodakarensis* chromosome. As shown below, histone, Alba, and TK0471/TrmBL2 are the major DNA-binding proteins on the *T. kodakarensis* chromosome that create different structures on DNA. The fact that similar structures were observed in the chromosome fibers released from the cell, as well as in isolated or reconstituted chromatin, supports the existence of these structures in the cell.

Consistent with the fluorescence imaging result was the observation that the chromosome released from cells in the stationary phase appeared in a condensed form under AFM. Individual fibrous structures could not be detected from condensed stationary phase chromosomes (Figure 1A).

A putative helix-turn-helix protein is one of the major chromosomal proteins in *T. kodakarensis*

Cell-free extract of *T. kodakarensis* was prepared through gentle treatment of the cells with a detergent, followed by a low-speed centrifugation through sorbitol (Matsunaga *et al.*, 2001). Chromatin was isolated as an insoluble pellet (hereafter referred to as the chromatin fraction). Separation of the protein component of the chromatin fraction through SDS–PAGE followed by Coomassie staining resulted in 20–30 visible bands, ranging from 6 to 120 kDa in size (Figure 2, A and B). The chromatin fraction exhibited a pattern of proteins different from that exhibited by the whole cell lysate (Figure 2A). In addition, Western blotting experiments showed that membrane-bound Lon protease (Lon_{Tk}) (Fukui *et al.*, 2002) was not present in the chromatin fraction (Figure 2A), indicating that the chromatin fraction was not contaminated with cell membrane or membrane-bound proteins. The fact that the chromosomes could easily be isolated indicates an important difference between the *T. kodakarensis* chromosomes and the *Escherichia coli* nucleoid, which cannot be separated from the cell membrane (Zimmerman, 2006). There was a slight difference in protein composition between the chromatin fractions taken from cells in the log phase and those taken from cells in the stationary phase (Figure 2B).

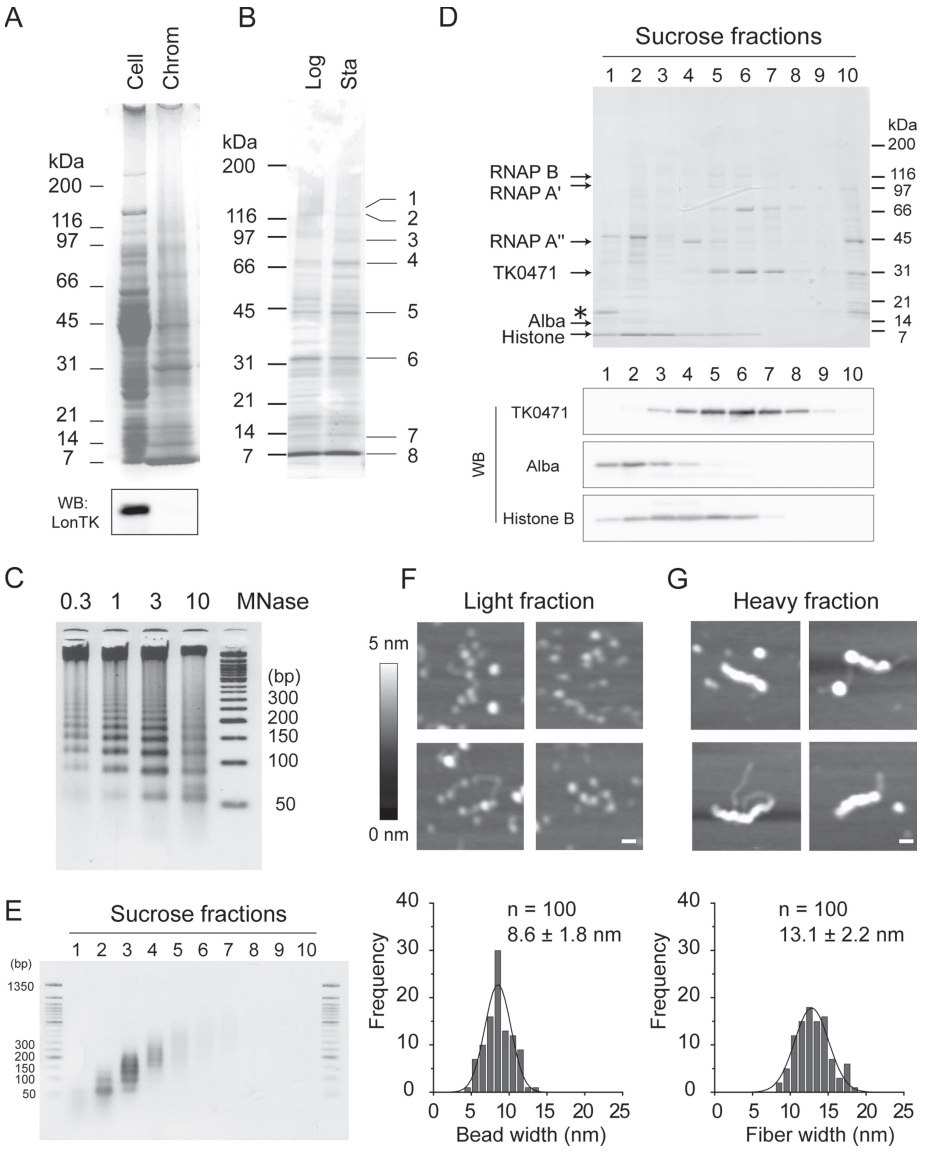


FIGURE 2: Separation of chromosome fragments with different structures and protein components. (A) Protein components of whole cell ("Cell") and of the chromatin fraction ("Chrom") of the KOD1 strain. Western blotting with an antiserum against the membrane-bound protease Lon_{TK} is shown at the bottom. (B) Protein components of chromatin fractions of KOD1 in the log and stationary phases. The protein bands analyzed through mass spectrometry are numbered on the right. The identified proteins were 1) predicted endonuclease-methyltransferase fusion protein (TK1460); 2) RNA polymerase (RNAP) subunit B (TK1083); 3) RNAP subunit A′ (TK1082); 4) acylamino acid-releasing enzyme (TK0752); 5) RNAP subunit A″ (TK1081) and glutamate dehydrogenase (TK1431); 6) TrmBL2 (TK0471) and fluctose-bisphosphate aldolase (TK0989); 7) Alba (TK0560); and 8) Histone A (TK1413), Histone B (TK2289), RPAP subunits N (TK1499), K (TK1498), and P (TK0616), and hypothetical proteins TK1040 and TK1270. Histones were estimated to constitute ~95% of band 8. (C) Chromatin fraction of KOD1 was mildly digested with increasing amounts of MNase (0.3, 1, 3, and 10 U/μl) and separated on 4% NuSieve GTG agarose gel in 1× TBE. (D) MNase-digested KOD1 chromatin fraction taken from stationary phase cells was separated through 5–20% sucrose density gradient sedimentation. The protein in each fraction was concentrated by means of trichloroacetic acid precipitation and separated on 5–20% SDS-polyacrylamide gel. The positions of known protein bands are indicated. The asterisk indicates RNase A. Western blotting with antisera against TK0471, Alba, or Histone B is shown at the bottom. (E) DNA was extracted and separated on 2.5% agarose gel in 0.5× TBE buffer. AFM images of chromosome fragments in (F) the light fraction (fraction 3) and (G) the heavy fraction (fraction 8). Scale bars, 50 nm. Histograms below the images indicate the diameter of each structure. The means ± SD of the distributions are indicated. The curves indicate the fitted Gaussian functions.

Protein bands of high intensity in the chromatin fraction of the KOD1 strain in the stationary phase were excised from the gel and analyzed through mass spectrometry. The identified proteins include DNA-binding proteins such as Histone A, Histone B, Alba, TK0471, and RNA polymerase (RNAP) subunits A′, A″, B, N, K, and P (Figure 2B). Micrococcal nuclease (MNase) digestion of the chromatin fraction resulted in a laddered pattern of DNA (~60, 90, 120, 150 base pairs, etc.) indicating the presence of a nucleosome structure in vivo (Figure 2C). Among the bands of DNA-binding proteins, the ~6-kDa band, which contains histones, had the highest intensity, but that of the ~32-kDa band was also high; the ~32-kDa band contains TK0471, a protein that has a putative helix-turn-helix (HTH) DNA-binding motif at the N-terminus.

The molecular ratios of Alba and TK0471 relative to histone were estimated to be 0.19 ± 0.06 and 0.10 ± 0.02, respectively, by dividing the intensity of each band in SDS–PAGE by the molecular weight of each protein (Supplementary Figure S1). In addition to these DNA-binding proteins, proteins without an obvious DNA-binding function were identified (see Figure 2B caption).

Each chromosomal protein forms a different structure on double-stranded DNA

The chromatin fraction of the *T. kodakarensis* strain KOD1 was partially digested with MNase and then subjected to 5–20% sucrose density gradient sedimentation in the expectation that the various structural elements would be separated according to their protein components and densities. Histones were concentrated in low- to middle-density sucrose fractions. TK0471 was concentrated in middle- to high-density fractions, and RNAP subunits were concentrated in middle-density fractions; this was observed through SDS–PAGE followed by Coomassie staining and Western blotting (Figure 2D). The distribution of Alba was observed in low-density fractions, partially overlapping the distribution of histone (Figure 2D).

The length of the DNA that was abundant in each fraction was determined through agarose gel electrophoresis (Figure 2E). The light fractions were enriched with small DNA fragments (60–300 bp). A characteristic laddered pattern (~60, 90, 120 bp, etc.) was observed in the light fractions, implying that the nucleosome structures were abundant in these fractions (Bailey *et al.*, 2000). The TK0471-rich heavy fractions were enriched with longer DNA fragments (200–1000 base pairs), although much less DNA was present in this fraction than in the lighter fractions (Figure 2E).

The structures of the chromosome fragments that accumulated in the light and heavy fractions were analyzed through AFM. Chromosome fragments in the light fraction (fraction 3) exhibited mainly a beads-on-a-string structure. The diameter of the beads was 8.6 ± 1.8 nm (mean ± SD, n = 100) (Figure 2F). On the basis of this observation, as well as on the histone accumulation and the laddered pattern of the DNA in this fraction (Figure 2, D and E), we concluded that the light fraction contains a high concentration of the archaeal nucleosome structure. In contrast, the heavy fraction (fraction 8) contained mainly chromosome fragments in a thick fibrous structure, the diameter of which was 13.1 ± 2.2 nm (mean ± SD, n = 100) (Figure 2G). Both the beads-on-a-string structure and the fibrous structure were detected in both the log phase and the stationary phase in KUW1, another wild-type strain (see below and Figure 5A later in the paper).

Next we tested the ability of histone, Alba, and TK0471 proteins to reconstitute different structures in vitro. Recombinant Histone A (HpkA), Alba, or TK0471 proteins were expressed in *E. coli* and purified. Incubation of each protein with a linear 3-kbp DNA altered the mobility of the DNA through an agarose gel (Figure 3, A–C). The structures of the resulting DNA–protein complexes were examined through AFM (Figure 3, D–G). The diameter and contour length of the 3-kbp dsDNA molecule used in this experiment were 1.83 ± 0.37 nm (n = 100) and 1003 ± 30 nm (mean ± SD, n = 100), respectively (Figure 3, H and L), which correspond to the theoretical diameter (2 nm) and the length (0.34 nm/bp) of B-form DNA. Beads-on-a-string structures were formed when Histone A (HpkA) was mixed with linear 3-kbp DNA (Figure 3E). The diameter of the beads was 8.9 ± 1.5 nm (n = 150) (Figure 3I). The contour length of the histone-bound DNA was 607 ± 74 nm (n = 100) (Figure 3M), significantly shorter than that of the naked DNA ($P < 0.01$, Tukey's HSD test). Taking into account the number of nucleosomes formed on each DNA molecule (10.1 ± 1.5, n = 54), we concluded that 113 ± 16 base pairs (38.6 ± 5.3 nm, n = 54) of DNA were incorporated into a single nucleosome on average. Incubation of TK0471 with DNA gave the genetic material a thick fibrous structure (Figure 3F) with a diameter of 14.2 ± 2.1 nm (n = 150) (Figure 3J), similar to that of the fibers observed in the TK0471-rich heavy fractions (Figure 2G). Incubation of Alba with DNA gave the genetic material a fibrous structure with a diameter of 10.3 ± 1.1 nm (n = 150) (Figure 3, G and K), thinner than that of the fibers formed in the presence of TK0471. The individual fibrous structure made with Alba, but not the contour DNA length, was shorter compared to that made with TK0471 (Figure 3, F and G). In contrast to histone, neither TK0471 (994 ± 31 nm, n = 100) nor Alba (999 ± 34 nm, n = 100) significantly reduced DNA contour length ($P = 0.52$ and 0.92, respectively, Tukey's HSD test) (Figure 3, N and O).

Histone and TK0471 localize to both coding and noncoding regions

Localization of each chromosomal structure on the genome was determined through massively parallel sequencing of DNA accumulated in the light and heavy fractions in the log phase. Fraction 3 was chosen as the histone-rich light fraction. Although in the light fractions the distribution of Alba and histone overlap, we expect the majority of DNA in this fraction to represent histone-bound DNA, based on the relative abundance of histone and the laddered pattern of DNA characteristic of nucleosome formation (Figure 2, D and E). Fraction 8 was chosen as the TK0471-rich heavy fraction. The complete profiles of the entire genomic regions that were concentrated in each fraction are shown in Supplementary Figure S2, and the nucleotide positions of the genomic sequences that were present at more than twice their normal concentration in each fraction are listed in Supplementary Table S1.

Peaks corresponding to the genomic regions that were concentrated in the light and heavy fractions were located throughout the genome and encompassed both coding and intergenic regions (Figure 4A). Among the 397 genomic regions that were concentrated more than twofold in the light fraction, 379 (95.5%) were located within the coding regions [open reading frames (ORFs) and RNA coding genes] and 18 (4.5%) were located in the intergenic regions. Among the 449 genomic regions that were concentrated more than twofold in the heavy fraction, 407 (90.6%) were located within the coding regions and 42 (9.4%) were located in the intergenic regions. Ninety-two percent of the *T. kodakarensis* genome has been predicted to consist of coding regions (Fukui *et al.*, 2005). The distributions of the peaks in each fraction were not significantly biased toward the coding or intergenic parts of the genome compared to a random distribution of peaks along the genome. In addition, the distributions of peaks in the two fractions were not significantly different from each other (light fraction versus random distribution on the genome, $P = 0.028$; heavy fraction versus random distribution on the genome, $P = 0.772$; light fraction

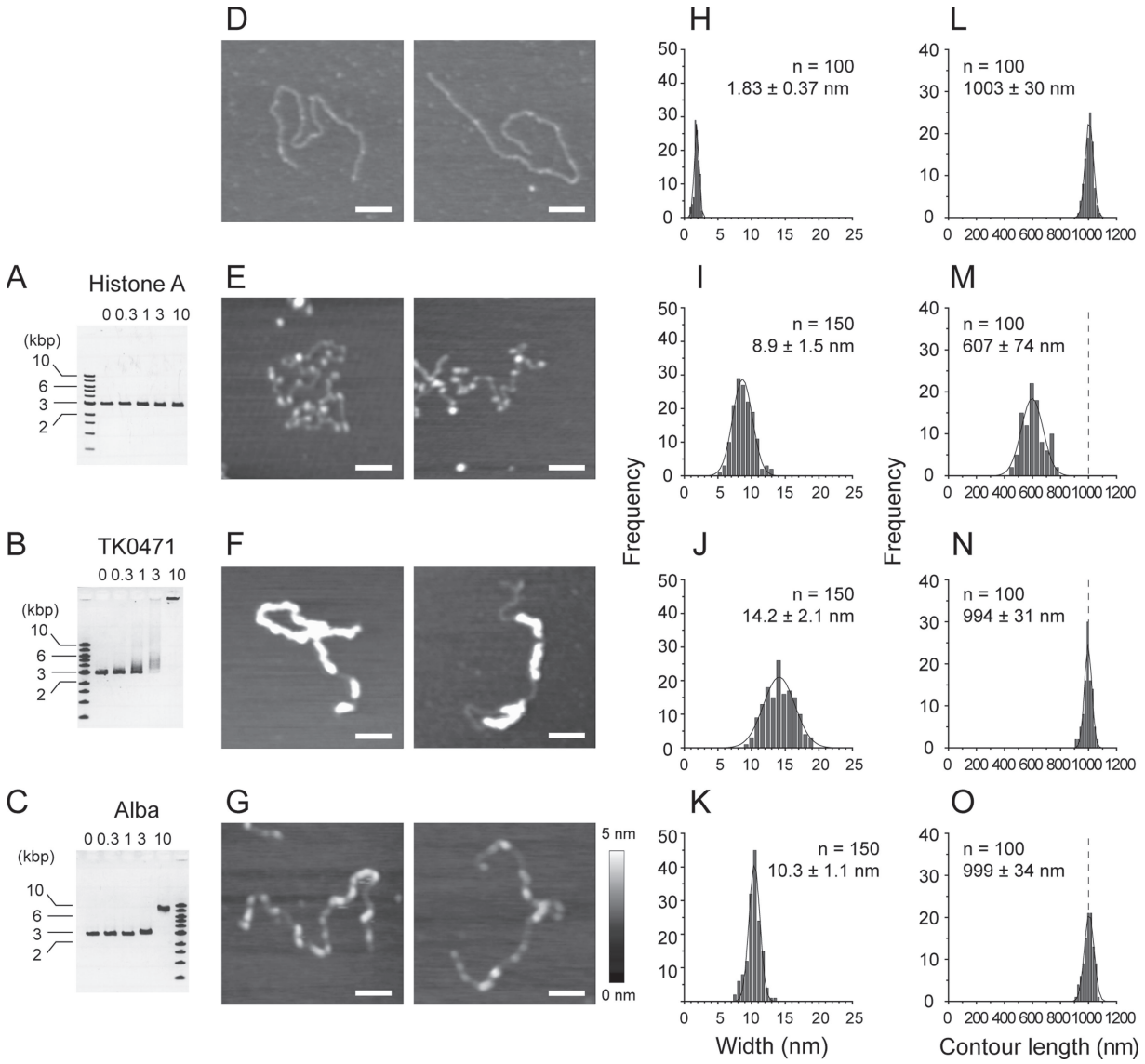


FIGURE 3: Recombinant chromatin proteins form distinct structures on dsDNA in vitro. (A-C) Mobility shift of DNA upon binding of recombinant chromatin proteins. One hundred nanograms of linear 3-kbp DNA (pBluescript II) was incubated with (A) Histone A, (B) TK0471, or (C) Alba. The protein-to-DNA ratio (wt/wt) was 0, 0.3, 1, 3, or 10 as indicated in each lane. (D–G) AFM images of (D) 3-kbp linear DNA, (E) 3-kbp linear DNA incubated with recombinant Histone A, (F) TK0471, or (G) Alba at a protein-to-DNA ratio of 10:1 (wt/wt). Scale bars, 100 nm. (H–K) Diameters of (H) naked dsDNA and those of the structures formed in the DNA by (I) Histone A, (J) TK0471, and (K) Alba are shown in histograms. (L) Contour length of the 3-kbp DNA molecule without protein and lengths of the molecule when bound with recombinant (M) Histone A, (N) TK0471, and (O) Alba are indicated in histograms. The means ± SD of the distributions are indicated. The curves indicate the fitted Gaussian functions.

versus heavy fraction, P = 0.021; Fisher's exact test with Bonferroni correction).

Our statistical analysis indicates that nucleosomes and TK0471 fibers tend not to colocalize in the log phase (see *Materials and Methods*). As an example, a profile of the DNA sequences that were concentrated in the light and heavy fractions in a 50-kb region including both intergenic and coding regions is shown in Figure 4B. Genomic sequences corresponding to the genes encoding DNA-directed RNA polymerase subunits A″ (TK1081), A′ (TK1082), and B (TK1083) (Figure 4C) as well as 16S and 23S ribosomal RNA genes (Tkr05 and Tkr06) (Figure 4D), were particularly highly concentrated in the TK0471-rich heavy fraction.

The DNA sequence accumulated in the light and the heavy fractions in the stationary phase was also determined. In this case, DNA in each fraction was purified and cloned into a plasmid vector, then subjected to conventional Sanger sequencing. The entire list of genomic regions cloned from each fraction is shown in Supplementary Table S1. Statistical analysis indicates that the sequence of DNA in the light fraction tends to differ in the log and stationary phases, but some portion of the DNA in the heavy fraction stays at the same position in both the log and stationary phases (see *Materials and Methods* for details).

Deletion of TK0471 gene leads to a loss of thick fibrous structure

A deletion strain of the TK0471 gene was constructed using a gene deletion system previously developed for this organism (Sato *et al.*, 2005). *T. kodakarensis* KUW1 strain (Δ*pyrF*, Δ*trpE*) was used as the

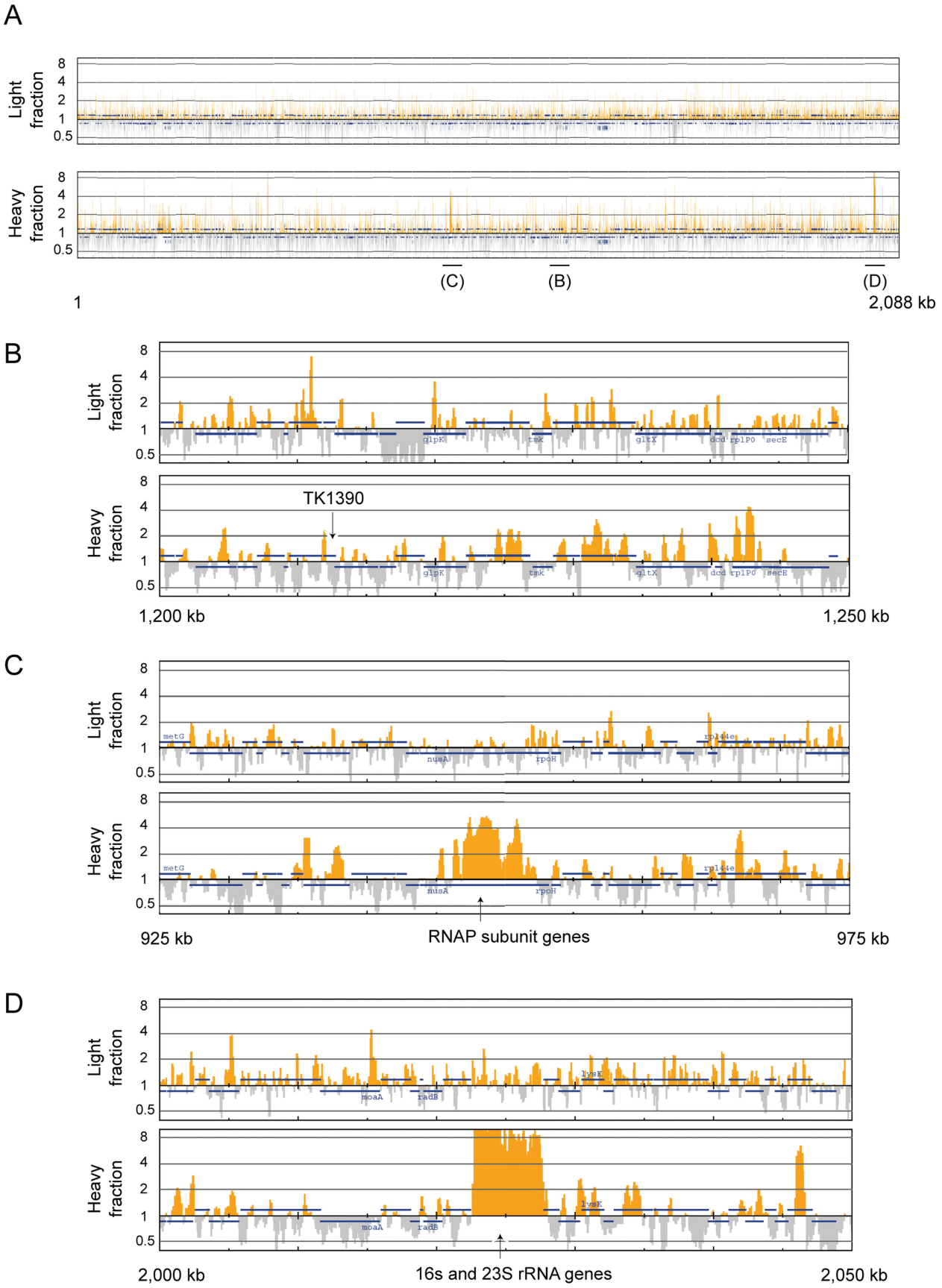


FIGURE 4: Profiles of the DNA regions found at high concentrations in the light and heavy fractions. The *x* axis indicates position on the genome. The blue bars above and below the baseline indicate predicted ORFs on the plus and minus strands, respectively. Note that a single bar can contain multiple overlapping ORFs. The *y* axis indicates the fold enrichment value of each sequence compared to that of the genomic DNA control. In each case, the profile of the light

host strain to allow the use of the *pyrF* gene as a selectable marker (Supplementary Figure S3A). KUW1 did not differ from KOD1 in terms of growth characteristics, protein composition in the chromatin fraction, or chromosome fiber structure. Successful deletion of the TK0471 gene was confirmed through PCR and Southern blotting (Supplementary Figure S3B), and the deletion strain (Δ*pyrF*, Δ*trpE*, ΔTK0471) was designated KCP1. The KCP1 strain exhibited a slight growth delay compared to KUW1 (Supplementary Figure S3C). When the chromatin fraction of KCP1 in the log phase was subjected to MNase digestion and sucrose density gradient sedimentation, the fibrous structures observed in the heavy fraction in wild-type KOD1 or KUW1 strains were not detected in any of the fractions (Figures 2G and 5A), supporting the conclusion that TK0471 is the protein component of these thick fibrous structures. Chromosome condensation was observed in the KCP1 strain in the stationary phase, as in the wild-type KOD1 and KUW1 strains (Figures 1A and 5B), indicating that TK0471 is not essential for condensation.

The chromatin fraction of the KCP1 strain was more sensitive to MNase digestion than that of the wild-type KUW1 strain (Figure 5C). Digestion with 0.3 U of MNase resulted in an almost complete removal of high-molecular-weight genomic DNA from the chromatin fraction of KCP1, whereas high-molecular-weight genomic DNA was left partially undigested in the chromatin fraction of KUW1 (Figure 5C). MNase digestion of the chromatin fraction of the KCP1 cells still resulted in the laddered pattern of DNA in the low-molecular-weight region that is characteristic of nucleosome formation (Figure 5C).

A protein band with an apparent molecular mass of ~35 kDa accumulated in the chromatin fraction of the KCP1 strain, although there was no obvious change in the patterns of other proteins compared to their patterns in the wild-type strains (KOD1 and KUW1) (Figure 5D). Mass spectrometric analysis revealed that the ~35-kDa band contained TK1959 and TK1961 proteins, which are archaeal homologues of eukaryotic replication protein A (RPA) complex subunits (Komori and Ishino, 2001).

TK0471 deletion results in up-regulation of specific genes in log phase

cDNAs generated from the RNA preparations of the wild-type (KUW1) and ΔTK0471 (KCP1) cells grown to the log phase in ASW-YT-S^0 medium were labeled with Cy3 and Cy5, respectively, and hybridized to the *T. kodakarensis* genomic array, which has spots that represent all 2306 predicted ORFs identified in the *T. kodakarensis* genome (Fukui *et al.*, 2005). The entire result is shown in Supplementary Table S2 and summarized in Figure 6A. In the KCP1 strain, the signal intensities of 151 ORFs on the genome (6.5% of the total) were elevated more than twofold compared to those in the wild-type KUW1 strain. However, 41 ORFs (1.8% of the total) were decreased more than twofold. The number of ORFs that did not exhibit a significant (i.e., twofold or greater) alteration was 1974.

The alteration in transcript level had a strong correlation with the presence of TK0471 protein on the promoter region in the log phase, which was determined by the sequencing of DNA in the heavy fraction (Figure 4 and Supplementary Table S1). The signal intensity ratio ($\log_2$[KCP1/KUW1]) of the ORFs that exhibited TK0471 enrichment in the promoter region (defined here as −200 to +50 relative to start codon) was 1.76 ± 3.27 (median ± quartile range, n = 90; Figure 6B and Supplementary Table S3). This was significantly higher than that of the ORFs with TK0471 enrichment in the coding region (−0.12 ± 0.59, n = 293; Figure 6C) or that of the ORFs that did not have significant enrichment of TK0471 (−0.16 ± 0.56, n = 1776; Figure 6D) ($P < 6.6 \times 10^{-16}$ in both cases, Wilcoxon rank sum test with Bonferroni correction). The signal intensity ratios of individual ORFs, together with the list of TK0471 enrichment, can be found in Supplementary Table S2.

Transcript levels of ORFs in the KCP1 strain in the stationary phase were also determined (Supplementary Table S2). Results were obtained for 1253 out of 2306 ORFs. The signals of the other ORFs could not be detected, probably due to degradation of the RNA in the stationary phase. Among the 1253 ORFs for which transcript levels were determined, 192 exhibited elevated (greater than twofold) signal intensities compared to those in the wild-type KUW1 strain, whereas 224 exhibited decreased (greater than twofold) signal intensities. The number of ORFs that did not exhibit significant (i.e., twofold or greater) difference between KCP1 and KUW1 was 837.

In contrast to the log phase, TK0471 enrichment to the promoter did not correlate with alteration of the transcript level in the ΔTK0471 strain (KCP1) in the stationary phase. The signal intensity ratio ($\log_2$[KCP1/KUW1]) of the ORFs that exhibited TK0471 enrichment in the promoter region (−200 to +50 relative to start codon) was 0.38 ± 0.84 (median ± quartile range, n = 11), which is not significantly different from that of the ORFs with TK0471 enrichment in the coding region (−0.41 ± 1.07, n = 57) or from that of the ORFs without any TK0471 enrichment (−0.16 ± 1.34, n = 1185) (P = 0.53 and P = 0.11, respectively). The complete set of microarray results, together with the list of ORFs with TK0471 enrichment, is shown in Supplementary Table S2.

DISCUSSION

Our study had four major findings: 1) In addition to histone, a considerable amount of TK0471 protein exists on the chromosome of the hyperthermophilic archaeon *T. kodakarensis*. 2) TK0471 forms a thick fibrous structure distinct from the archaeal nucleosome array, providing heterogeneity in chromosome architecture. 3) TK0471 acts as an architectural protein for chromosomes as well as a global transcriptional repressor. 4) The localization of TK0471 changes dynamically depending on the growth stage of the cells.

Archaeal genome architecture

The results of our reconstitution experiment (Figure 3) suggest that, among the major DNA-binding proteins on the *T. kodakarensis* chromosome, histone is the only one that contributes to compaction by wrapping genomic DNA. Digestion of the isolated chromatin fraction with MNase resulted in a characteristic laddered pattern of DNA (~60, ~90, 120, 150 base pairs, etc.; Figure 2C). The ~60- and ~90-base pair DNA fragments correspond to the length of the fragment that makes direct contact with the histone fold of the archaeal histone tetramer and to the length of the fragment that fully circumvents the histone core, respectively (Pereira *et al.*, 1997; Pereira and Reeve, 1999). About 120 base pairs of DNA correspond to the length of DNA that is wrapped around a histone tetramer

fraction (fraction 3) is shown in the top panel, and that of the heavy fraction (fraction 8) in the bottom panel. (A) Whole genome profile of the sequences highly concentrated in the light and heavy fractions. (B–D) Magnified profile of the genomic region including (B) the TK1390 gene, (C) RNAP subunit genes, and (D) 16S and 23S rRNA genes. See also Supplementary Figure S2 and Supplementary Table S1.

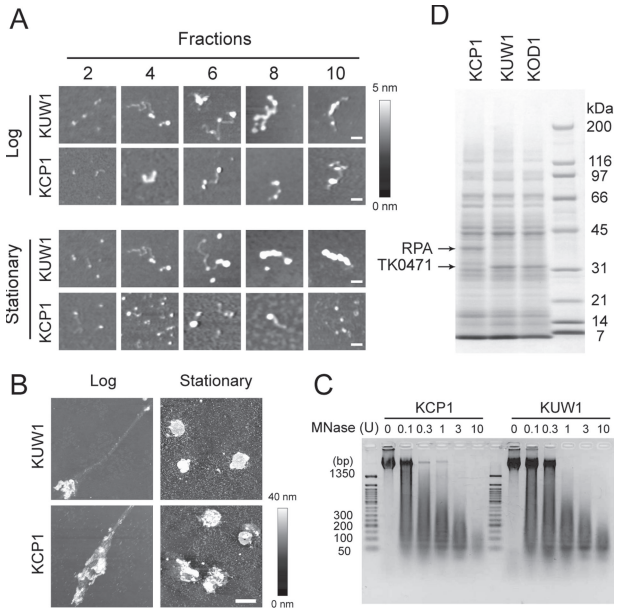


FIGURE 5: Analysis of the ΔTK0471 strain. (A) MNase-digested chromatin fractions of KUW1 and KCP1 strain (each in both log phase and stationary phase) were separated through 5–20% sucrose density gradient sedimentation. The thick fibrous structures that are characteristic of TK0471 were not observed in the KCP1 strain. Scale bars, 50 nm. (B) Chromosome condensation in the stationary phase occurs in the KCP1 strain. Scale bar, 2 μm. (C) Sensitivity of the chromatin fractions of (left) KCP1 and (right) KUW1 to MNase digestion. Chromatin fractions were incubated with increasing amounts of MNase (0, 0.1, 0.3, 1, and 10 U) at 37°C for 1 h. DNA was purified and separated on 2.5% agarose in 0.5× TBE buffer. (D) Proteins in the chromatin fractions of the KOD1, KUW1, and KCP1 strains were compared by means of SDS–PAGE on a 5–20% gradient gel. In the KCP1 strain, the ~32-kDa protein band disappeared and a band of ~35 kDa appeared. The ~35-kDa band was identified through mass spectrometry as consisting of RPAs. See also Supplementary Figure S3.

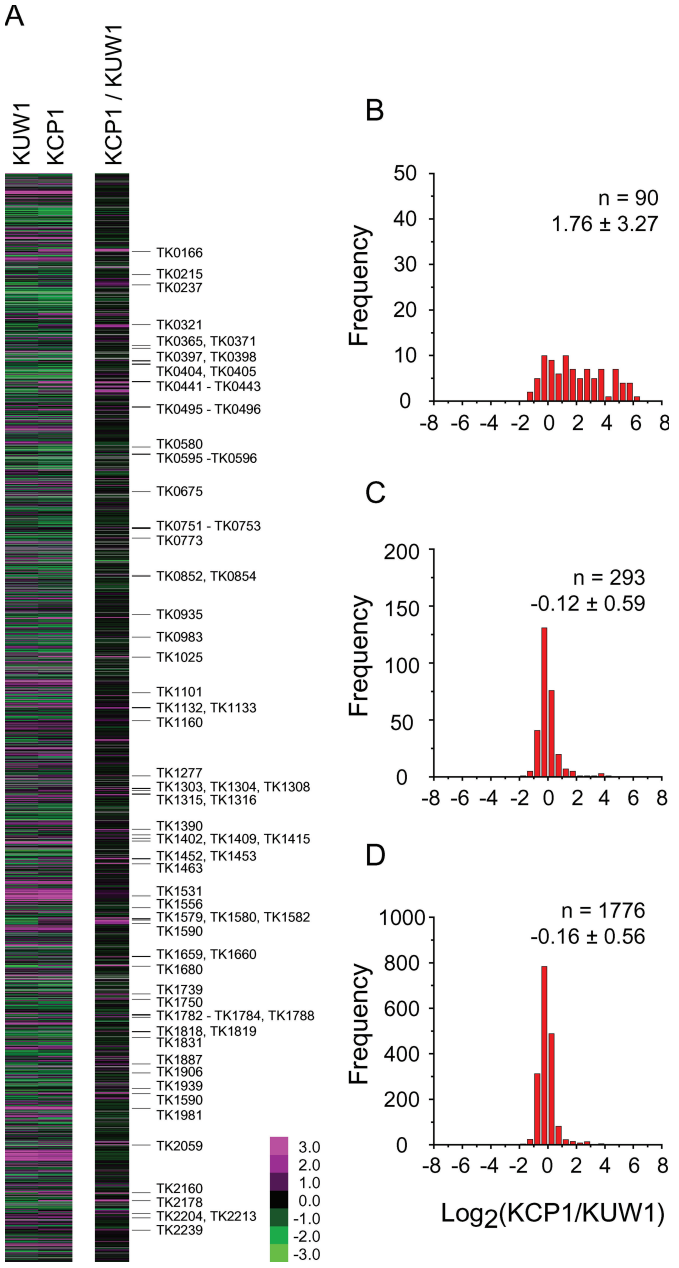
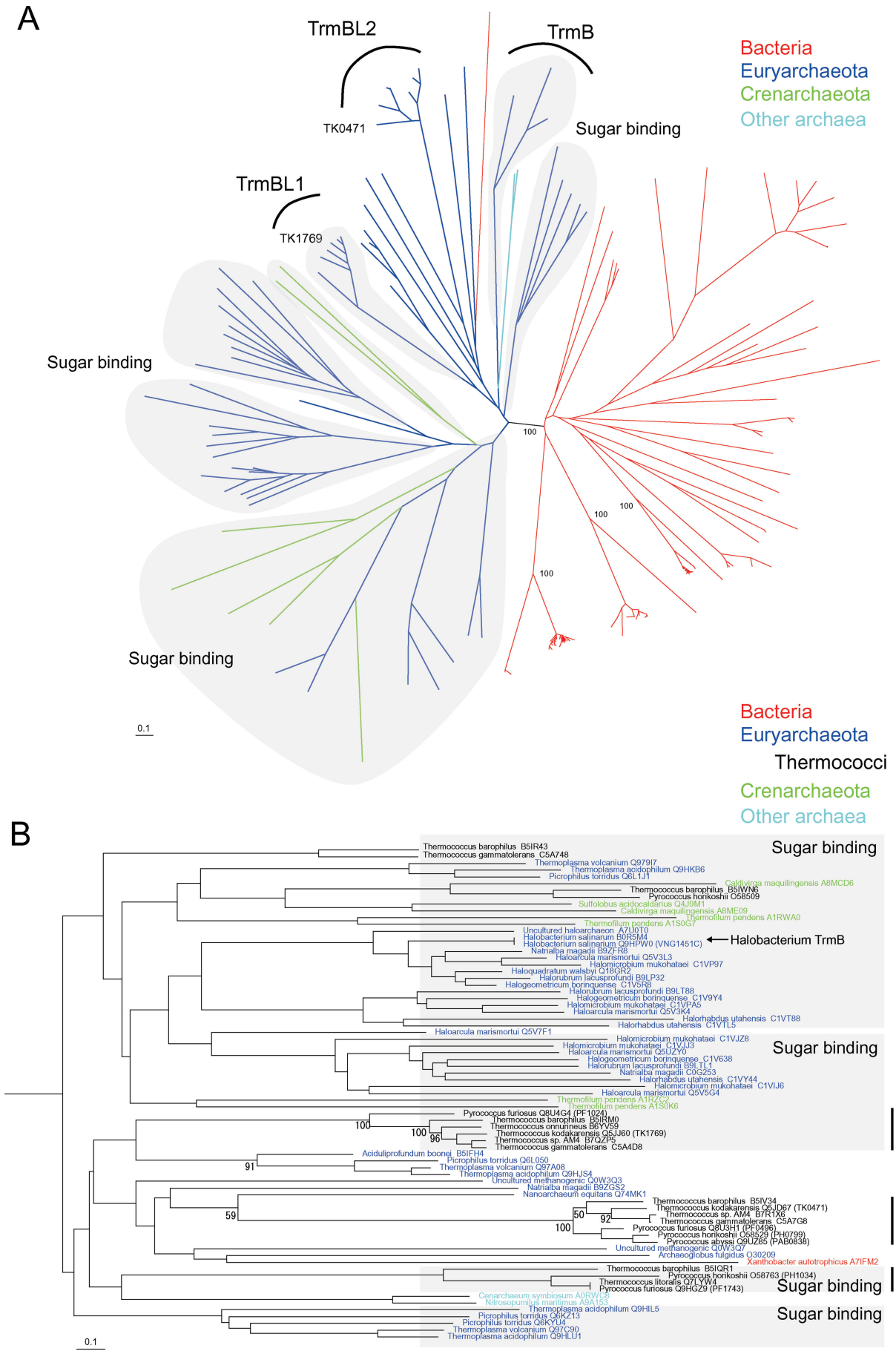


FIGURE 6: Transcriptome analysis in *T. kodakarensis* KUW1 and KCP1 strains in the log phase. (A) Transcript levels of each predicted ORF in KUW1 and KCP1, as well as the "KCP1/KUW1" ratio, are indicated by color. Magenta represents relatively high signal intensity while green indicates relatively low signal intensity. In the right column, magenta indicates a high signal intensity ratio, while green indicates a low signal intensity ratio ($\log_2$[KCP1/KUW1]). ORFs with no data are indicated in gray. ORFs that exhibited TK0471 enrichment in the promoter region are indicated on the right. (B–D) Signal intensity ratios ($\log_2$[KCP1/KUW1]) of ORFs, classified according to whether TK0471 is (B) bound to a promoter region (−200 to +50 relative to start codon), (C) bound to a coding region, or (D) not bound to any region, are indicated by histograms. The median ± quartile range of each sample is indicated. Promoter binding of TK0471 correlates with transcript elevation in the ΔTK0471 strain (KCP1). See also Supplementary Tables S2 and S3.

(Tomschik *et al.*, 2001). The longer DNA fragments (~150 base pairs and larger) may correspond to the lengths of fragments that are in contact with more than one histone tetramer.

In the *T. kodakarensis* genome, the TK0471 protein forms a thick fibrous structure (Figures 2G and 3F). Even in the histone-encoding euryarchaeal species, histone-based DNA wrapping is not the only mechanism that plays a role in genome organization. Indeed earlier electron microscopic studies of the chromosome of the halophilic euryarchaeon *Halobacterium salinarium* have shown that nucleosome-free regions exist on the chromosome (Shioda *et al.*, 1989; Takayanagi *et al.*, 1992). On the basis of its distribution throughout the chromosome, we propose that TK0471/TrmBL2 is a novel chromosomal protein that all species in the order *Thermococcales* possess (see below). In support of this notion, TrmBL2 of *Pyrococcus furiosus* (PF0496) was shown to be one of the proteins in the cell extract that can bind to DNA in a sequence-independent manner in vitro (Lipscomb *et al.*, 2009).

We show here that the protein Alba of *T. kodakarensis* can form, in vitro, a short fibrous structure on dsDNA (Figure 3G). This structure is similar in diameter to that of the "helical protein fiber" formed in the presence of the Alba of the *Sulfolobus* species, as observed by electron microscopy (Lurz *et al.*, 1986) or modeled from the crystal structure (Wardleworth *et al.*, 2002). However, such a structure was not identified in the light fractions of the chromatin isolated from the cell (Figure 2F). This is likely due to the small amount of Alba relative to histone. Because the distribution of Alba overlaps with and cannot completely be separated from that of histone in our


A
TrmBL2
TrmB
TK0471
Sugar binding
TrmBL1
TK1769
Sugar binding
Sugar binding
100
100
100
100
0.1
Bacteria
Euryarchaeota
Crenarchaeota
Other archaea
Bacteria
Euryarchaeota
Thermococci
Crenarchaeota
Other archaea
B
Thermococcus barophilus B5IR43
Thermococcus gammatolerans C5A748
Thermoplasma volcanium Q979I7
Thermoplasma acidophilum Q9HKB6
Picrophilus torridus Q6L1J1
Caldivirga maquilingensis A8MCD6
Thermococcus barophilus B5IWN6
Pyrococcus horikoshii O58509
Sulfolobus acidocaldarius Q4J9M1
Caldivirga maquilingensis A8ME09
Thermofilum pendens A1RWA0
Thermofilum pendens A1S0G7
Uncultured haloarchaeon A7U0T0
Halobacterium salinarum B0R5M4
Halobacterium salinarium Q9HPW0 (VNG1451C)
Halobacterium TrmB
Natrialba magadii B9ZFR8
Haloarcula marismortui Q5V3L3
Halomicrobium mukohataei C1VP97
Haloquadratum walsbyi Q18GR2
Halorubrum lacusprofundi B9LP32
Halogeometricum borinquense C1V5R8
Halorubrum lacusprofundi B9LT88
Halogeometricum borinquense C1V9Y4
Halomicrobium mukohataei C1VPA5
Haloarcula marismortui Q5V3K4
Halorhabdus utahensis C1VT88
Halorhabdus utahensis C1VTL5
Haloarcula marismortui Q5V7F1
Halomicrobium mukohataei C1VJZ8
Halomicrobium mukohataei C1VJJ3
Haloarcula marismortui Q5UZY0
Halogeometricum borinquense C1V638
Halorubrum lacusprofundi B9LTJ1
Natrialba magadii C0GZ53
Halorhabdus utahensis C1VY44
Halomicrobium mukohataei C1VIJ6
Haloarcula marismortui Q5V5G4
Thermofilum pendens A1RZC2
Thermofilum pendens A1S0K6
Pyrococcus furiosus Q8U4G4 (PF1024)
Thermococcus barophilus B5IRM0
Thermococcus onnurineus B6YV59
Thermococcus kodakarensis Q5JJ60 (TK1769)
Thermococcus sp. AM4 B7QZP5
Thermococcus gammatolerans C5A4D8
Aciduliprofundum boonei B5IFH4
Picrophilus torridus Q6L050
Thermoplasma volcanium Q97A08
Thermoplasma acidophilum Q9HJS4
Uncultured methanogenic Q0W3Q3
Natrialba magadii B9ZGS2
Nanoarchaeum equitans Q74MK1
Thermococcus barophilus B5IV34
Thermococcus kodakarensis Q5JD67 (TK0471)
Thermococcus sp. AM4 B7R1X6
Thermococcus gammatolerans C5A7G8
Pyrococcus furiosus Q8U3H1 (PF0496)
Pyrococcus horikoshii O58529 (PH0799)
Pyrococcus abyssi Q9UZ85 (PAB0838)
Uncultured methanogenic Q0W3Q7
Archaeoglobus fulgidus O30209
Xanthobacter autotrophicus A7IFM2
Thermococcus barophilus B5IQR1
Pyrococcus horikoshii O58763 (PH1034)
Thermococcus litoralis Q7LYW4
Pyrococcus furiosus Q9HGZ9 (PF1743)
Cenarchaeum symbiosum A0RWC8
Nitrosopumilus maritimus A9A153
Thermoplasma acidophilum Q9HIL5
Picrophilus torridus Q6KZ13
Picrophilus torridus Q6KYU4
Thermoplasma volcanium Q97C90
Thermoplasma acidophilum Q9HLU1
100
100
96
91
59
50
92
100
Sugar binding
Sugar binding
Sugar binding
TrmBL1
TrmBL2
TrmB
0.1

experimental system, we did not determine the localization of Alba along the genome. A chromatin immunoprecipitation approach would be necessary in the future to identify the precise localization of Alba. It has also been shown that Alba can hold two DNA duplex together in vitro (Lurz *et al.*, 1986; Jelinska *et al.*, 2005). Although we could not detect such a structure under the condition that we used (for instance, we used a linear instead of circular DNA molecule), it might occur in vivo.

The *T. kodakarensis* chromosome undergoes dynamic structural changes during the transition from the log phase to the stationary phase (Figure 1A). Because chromosome condensation still occurs in the ΔTK0471 strain (KCP1) (Figure 5B), TK0471 is not essential for chromosome condensation. The slight difference in the protein composition of the chromatin fraction between the log and stationary phases (Figure 2B) might be involved in the chromosome condensation in the stationary phase. It would be interesting to figure out in the future which of these proteins are responsible for the chromosome condensation in the stationary phase.

Physiological roles of TK0471/TrmBL2

Sequencing of DNA coenriched with TK0471 illustrates that this protein localizes both to the coding regions and to the intergenic regions in vivo (Figure 4). This mode of localization is similar to that of the bacterial nucleoid-associated proteins, such as IHF and H-NS (Grainger *et al.*, 2006) and several recently studied archaeal transcription factors and regulators (Koide *et al.*, 2009). A significant correlation between TK0471's binding to the promoter region and alteration of the transcript levels of downstream genes (Figure 6B) suggests that TK0471 directly binds to some promoter regions and represses their transcription, probably by blocking RNAP's access. The thick fibrous structure would, in this case, be formed by the binding of TK0471 to dsDNA, followed by the incorporation of more proteins into the adjacent region of the DNA through protein–protein interactions in a manner similar to that by which H-NS binds to DNA (Bouffartigues *et al.*, 2007; Lang *et al.*, 2007). A significant number of the ORFs that were identified as being under direct control of TK0471 are classified as hypothetical proteins by genome annotation (Fukui *et al.*, 2005). The exact cellular role of these proteins in archaeal life is not known; additional experiments are required for their characterization.

In contrast to TK0471's binding to promoters, which repress transcription, TK0471's binding to coding regions does not disrupt transcription elongation (Figure 6C). Because the microarray used in this study contains ~300 base pairs of DNA corresponding to the 3′ region of each predicted ORF, we were able to determine that RNAP transcribes throughout the entire coding sequence rather than halting in the middle. TK0471's binding to the coding region may play as-yet-unidentified roles in genome structure and function. Our results suggest that one of these roles could consist of shaping the genetic material into a structure that makes it less accessible to enzymes or other proteins (Figure 5C).

It is interesting to note that, although the transcript levels of more than 150 ORFs vary by more than twofold between the wild-type (KUW1) and the ΔTK0471 strain (KCP1), only RPAs highly accumulated in the chromatin (Figure 5D). The transcripts of RPA subunit genes (TK1959 and TK1961) were approximately four to six times more abundant in the ΔTK0471 strain (2.53 ± 0.21 and 2.16 ± 0.10, respectively, in $\log_2$ scale), which rank 60th~80th on the list of ORFs with high ΔTK0471/wild-type transcript ratio (Supplementary Table S2). The high transcript level of the RPA genes, together with their reported DNA-binding activity (Komori and Ishino, 2001), may explain the abundance of RPA in the chromatin of the ΔTK0471 strain. The questions of whether the RPA transcription is under direct control of TK0471, and what physiological significance it has, are unresolved at present.

Given that TK0471 localizes to specific promoter regions of the chromosome, the existence of a consensus recognition sequence was expected, but an extensive search consisting of sequence analysis of DNA concentrated with TK0471 has thus far identified no candidates for such a consensus recognition sequence. We also did not detect any significant difference in the bending propensity of the DNA sequence compared to the whole genome, using the trinucleotide flexibility parameter described previously (Brukner *et al.*, 1995; Fukue *et al.*, 2004). It can be speculated that TK0471 might localize specifically to relatively nucleosome-free regions without having its own specific recognition sequence. It has been shown that archaeal histone, like eukaryotic histone, prefers certain sequence motifs (nucleosome positioning signals) when binding to DNA (Bailey *et al.*, 2000).

TK0471/TrmBL2 forms a superfamily

TK0471 shares amino acid sequences homologous to the transcription factors TrmB (transcriptional regulator of mal operon) and Tgr/TrmBL1 (*Thermococcales* glycolytic regulator/TrmB-like 1) of *Thermococcales* and thus was previously named TrmBL2 (TrmB-like 2) (Kanai *et al.*, 2007; Lee *et al.*, 2007). A search of the Pfam database (Pfam 24.0, October 2009, http://pfam.sanger.ac.uk/) revealed that these proteins each consist of two domains, TrmB (Pfam ID: PF01978) and Regulator_TrmB (Pfam ID: PF11495). The TrmB domain is a member of the HTH DNA-binding domain and is widely distributed in many bacteria, both Gram-positive and Gram-negative, as well as in archaea, including most archaeal classes identified so far (Perez-Rueda and Janga, 2010; Supplementary Table S4). The TrmB domain is found in 40 archaeal and 99 bacterial genera (out of the total of 48 archaeal and 364 bacterial genera with at least one sequenced genome listed in the KEGG database as of April 2010; http://www.kegg.jp/). Among these TrmB domain-containing proteins, there are proteins that consist only of the TrmB domain as well as proteins that possess domains other than Regulator_TrmB. The variation in domain organization, together with the wide distribution of this domain, suggests that the TrmB domain appeared in the early stage of life history and that it forms a large superfamily with DNA-binding ability.

Proteins with the same domain organization as TK0471/TrmBL2 (i.e., both TrmB and Regulator_TrmB) are found in 18 archaeal and 19 bacterial genera (Figure 7, A and B, and Supplementary Table S4). Given the total number of genera with sequenced genomes in each domain, namely 48 archaeal and 364 bacterial, this type of

FIGURE 7: Phylogenetic analysis of proteins that have both TrmB and Regulator_TrmB domains. (A) Unrooted phylogenetic tree of proteins using TrmB and Regulatory_TrmB domains for alignment. (B) Magnification of the left half of the tree shown in (A). The UniProt accession number of each protein is indicated. The TrmB family protein of *Halobacterium salinarum NRC-1* (VNG1451C) (Schmid *et al.*, 2009) is indicated by an arrow. Proteins that have a sugar binding domain at the C-terminus of the Regulator_TrmB domain are shaded in gray. For instance, TK0471/TrmBL2 lacks the sugar binding domain. Scale bars indicate the Dayhoff distance among proteins.

protein appears to be more common in archaea than in bacteria. Within the archaea, it is found in 13 genera in Euryarchaeota, 3 genera in Crenarchaeota, and 1 genus each in Korarchaeota and Thaumarchaeota. Halobacteriales, Thermococcales, and Thermoplasmatales are the euryarchaeal orders that possess this type of protein most frequently (Supplementary Table S4). It might have emerged in one of these genera and spread by horizontal gene transfer to other archaeal and bacterial species. It is interesting that this type of protein exists in a limited number of species. Many other chromosomal proteins in prokaryotes are also limited in their gene distribution; for instance, H-NS is conserved only in Gram-negative Gammaproteobacteria, and SarA, which was proposed as a functional counterpart of H-NS (Fujimoto *et al.*, 2009), is present only in Gram-positive Firmicutes. These facts suggest that prokaryotic species in general have evolved their own chromosomal proteins to organize their chromosomes in conjunction with transcriptional regulation. Further comparative studies between archaea and bacteria would provide more insights into the specific function of this type of proteins.

MATERIALS AND METHODS

Strains and growth conditions

T. kodakarensis KOD1 was cultivated under anaerobic conditions at 85°C in the nutrient-rich medium ASW-YT, supplemented with 0.2% elemental sulfur (ASW-YT-S^0). The ASW-YT medium was composed of a 1.25-fold dilution of artificial seawater (Robb and Place, 1995) (0.8X ASW; 16 g/l NaCl, 2.4 g/l $MgCl_2{\cdot}6H_2O$, 4.8 g/l $MgSO_4{\cdot}7H_2O$, 0.8 g/l $(NH_4)_2SO_4$, 0.16 g/l $NaHCO_3$, 0.24 g/l $CaCl_2{\cdot}2H_2O$, 0.4 g/l KCl, 0.34 g/l KH_2PO_4, 40 mg/l NaBr, 16 mg/l $SrCl_2{\cdot}6H_2O$, 8 mg/l $Fe(NH_4)$ citrate), 5.0 g/l of yeast extract, and 5.0 g/l of tryptone. Resazurin was added at a concentration of 0.8 mg/l. Prior to inoculation, Na_2S was added to the medium until it became transparent. A fresh medium was inoculated with a preculture in the stationary phase and cells were harvested at log or stationary phase. Construction of the deletion strain is described in the Supplemental Materials.

On-substrate cell lysis of *T. kodakarensis* cells

KOD1 cells were harvested and centrifuged for 1 min at 10,000 rpm at 4°C. Cells were washed and resuspended in 0.8x ASW. Cell density was adjusted to OD_{660} = ~0.5. Cell suspension (50 µl) was applied to a cover glass (15 mm round; Matsunami Glass Ind., Kishiwada, Japan) and left for 10 min at 25°C. Liquid was blown off with nitrogen or difluoroethane (HFC-152a) gas. Cells on the cover glass were lysed through addition of Milli-Q water and incubation for 10 min at 25°C. Liquid was blown off with gas and samples were subjected to AFM for observation. For DAPI staining, lysed cells attached to a cover glass were stained by adding 0.8× ASW containing DAPI onto the cover glasses (18 × 18 mm; Matsunami). Intact cells were stained by adding DAPI to the culture.

Atomic force microscopy

Samples were observed under a Nanoscope IIIa or IV (Digital Instruments, Santa Barbara, CA) in tapping mode using an OMCL-AC160TS probe (Olympus, Tokyo, Japan). Either type E or type J scanners were used, depending on the size of the observed area. Images were analyzed using Nanoscope software. Widths of the chromatin fibers were calculated based on apparent bottom-to-bottom distance of particles, using the equation described previously for the same probe: S = 0.75W – 16.14, where W is the apparent measured bottom-to-bottom distance and S is the estimated actual width of samples (Ohniwa *et al.*, 2007).

Preparation of chromatin fraction

Chromatin fraction was obtained by the method described previously (Matsunaga *et al.*, 2001) with a slight modification. *T. kodakarensis* cells were harvested in the log phase or the stationary phase. Samples were washed once with 0.8x ASW. Cell suspensions (~10^{10} cells per milliliter) were disrupted by the addition of extraction buffer (25 mM HEPES [pH 7.0], 15 mM $MgCl_2$, 100 mM NaCl, 0.4 M Sorbitol, and 0.5% Triton X-100). After incubating for 10 min at 4°C, soluble proteins (supernatant) and a chromosomal DNA-enriched insoluble fraction (pellet) were separated by centrifugation at 14,000 *g* for 20 min. The insoluble chromatin fraction, containing DNA and chromatin-associated protein, was washed with the extraction buffer and frozen at −20°C until further use.

Mass spectrometry

Proteins in the chromatin fraction were separated through SDS–PAGE on a 5–20% gradient polyacrylamide gel (Bio-Rad Laboratories, Hercules, CA) after digestion with MNase (Worthington Biochemical, Lakewood, NJ). Protein bands were excised from the gel and digested in-gel with trypsin. The matrix-assisted laser desorption/ionization time-of-flight mass spectra measurements were performed on an AXIMA-CFR spectrometer (Shimadzu Biotech, Kyoto, Japan) for peptide map fingerprinting in the case of bands 1–7, or by liquid chromatography ion trap time-of-flight measurements for MS/MS ion search in the case of band 8 in Figure 2B. Matching peptides and proteins were run through a MASCOT analysis. The exponentially modified protein abundance index (emPAI) of each protein identified in band 8 was as follows: Histone A (8.46); Histone B (59.35); RNAP subunits N (1.34), K (0.58), and P (0.74); TK1040 (0.62); and TK1270 (0.54). The relative abundance of proteins in band 8 (Figure 2B) was determined from the molecular weight and emPAI of each protein as described previously (Ishihama *et al.*, 2005). The amount of Histone A and Histone B together was estimated to be ~95% (wt) of total protein in band 8 (~12% histone A and ~83% histone B) and the calculated amount of histone was adjusted accordingly.

Sucrose density gradient sedimentation of MNase-digested chromatin

Each chromatin fraction extracted from 30 to 60 ml of culture was digested with 1 U of MNase (Worthington Biochemical) in 100 µl of MNase buffer (20 mM Tris-HCl [pH 7.4], 5 mM NaCl, 25 mM $CaCl_2$) for 30 min at 37°C in the presence of RNase A (10 µg/µl). The reaction was stopped by adding EDTA to 50 mM. Digested chromatin was applied to 5 ml of 5–20% sucrose density gradient in sucrose gradient (SG) buffer (20 mM Tris-HCl [pH 7.4], 5 mM NaCl, 50 mM EDTA), made in 13 × 51 mm thinwall polyallomer or Ultra-Clear tubes (Beckman Coulter, Brea, CA) using the Gradient Master device (BioComp Instruments, Fredericton, New Brunswick, Canada). Centrifugation was performed at 28,000 *g* for 16–24 h at 4°C in a swinging-bucket MLS-50 rotor (Beckman). Ten fractions (500 µl each) were collected from the surface by means of careful pipetting. Extraction of DNA and protein from each fraction is described in the Supplemental Materials. For AFM observation of chromosome fragments, 50 µl of each fraction was passed through a Micro Bio-Spin Tris column (Bio-Rad) preequilibrated with SG buffer and applied onto a mica surface (pretreated with spermidine) at an appropriate concentration after fixation with 0.3% glutaraldehyde (for 30 min at 25°C).

Preparation of recombinant proteins

Plasmids for expression of *T. kodakarensis* Histone A (HpkA; TK1413) and Histone B (HpkB; TK2289) in *E. coli* were kindly provided by H. Higashibata and S. Fujiwara (Higashibata *et al.*, 1999). Genes encoding Alba (TK0560) and TK0471 were amplified by PCR using genomic DNA as a template and were inserted into pET vectors (Novagen, Madison, WI), with no additional amino acid sequence added to the native sequence. *E. coli* BL21-CodonPlus (DE3)-RIL competent cells (Stratagene, La Jolla, CA) were transformed with these plasmids and were cultured to the log phase. Expression of the proteins was induced by the addition of IPTG to a concentration of 0.1 mM. The cells were harvested after 24 h of culturing at 20°C. Cells were sonicated and incubated for 20 min at 85°C. After centrifugation, the supernatant was collected and subjected to purification by ion exchange and gel filtration chromatography.

Reconstitution of chromatin structures using recombinant proteins

For in vitro reconstitution, pBluescript II (Stratagene) plasmid was digested by the restriction enzyme *Hind*III. The linear DNA was mixed with each recombinant protein (in 10 mM Tris [pH 8.0], 200 mM NaCl) in DNA-to-protein ratios (wt:wt) of 1:0, 1:0.3, 1:1, 1:3, or 1:10 and incubated for 20 min at 50°C. In our electrophoretic mobility shift assay, the protein–DNA complexes were run through 0.8% agarose gel in 1× TAE buffer. For AFM observation, the protein–DNA complexes were diluted to 1:20 in AFM fixation buffer (10 mM Tris [pH 8.0], 5 mM NaCl, 0.3% glutaraldehyde). After incubation for 30 min at 25°C, 10 µl of the mixture (equivalent to ~2.5 ng of DNA) was deposited on mica that had been pretreated with 10 mM spermidine for 10 min. After 10 min of incubation, the mica was washed with 1 ml of pure water and dried with nitrogen gas. AFM observation and analysis were performed as described above.

DNA sequencing

DNA in the light fraction (third fraction from the top) and the heavy fraction (eighth fraction) in the 5–20% sucrose density gradient sedimentation of MNase-digested chromatin fraction of the KOD1 strain in the log phase was purified through phenol/chloroform extraction and ethanol precipitation, following RNase A and Proteinase K treatment of each fraction. Purified DNA was sequenced using a SOLiD 3 sequencer (Applied Biosystems, Foster City, CA). Purified KOD1 genomic DNA was also sequenced as a control. The "fold enrichment" value was determined by dividing the degree of sequence enrichment in the sample with the degree of enrichment in the control sample. Peaks were defined as the middle nucleotide positions of genomic regions enriched more than twofold in a particular fraction. DNA in the light and heavy fractions in the stationary phase was purified, treated with DNA Polymerase I Klenow fragment (New England Biolabs, Ipswich, MA), cloned into a plasmid pBluescript II (Stratagene), and then sequenced with the Sanger method. Peaks were defined as the middle nucleotide positions of the cloned regions. The UCSC Archaeal genome Browser (http://archaea.ucsc.edu/) was used in part to analyze the data (Schneider *et al.*, 2006).

Microarray analysis

T. kodakarensis KUW1 (wild type) and KCP1 (ΔTK0471) strains were cultivated at 85°C in ASW-YT-S^0 media. Cells were harvested in the early log phase ($OD_{660} = 0.1$) and total RNA was extracted using the RNeasy Midi kit (Qiagen, Hilden, Germany). The microarray plate used in this study (Array Tko2) was manufactured at Takara Bio (Otsu, Japan) and covers all 2306 ORFs that have been predicted to exist in the *T. kodakarensis* genome. DNA fragments about 300 base pairs in length, corresponding to the 3′-terminal regions of each coding region, were spotted on the glass plate. Because two identical sets (left and right) were loaded onto each plate, two sets of data were obtained from each microarray plate. Average signal ratios and standard deviations are shown in our data files. Java TreeView was used to visualize the microarray data (Saldanha, 2004).

Statistical analysis

Analyses were performed using the statistical package R (http://www.r-project.org/) (R Development Core Team, 2010).

To evaluate the degree of overlap between the DNA sequences that were concentrated in each of the various fractions in the sucrose density gradient sedimentation experiment, a bimodal test with a Bonferroni correction was performed. In the log phase, the rate of overlap in nucleotide positions between the histone-rich light and TK0471-rich heavy fractions (7.3%) was significantly lower than the rate that would have been expected from a random distribution of the same number of nucleotide positions (7.7%) ($P = 0.0015$).

The nucleotide positions between the genomic regions that were concentrated in the light fraction in the stationary phase had a significantly lower rate of overlap (2.5%) with the light fraction in the log phase ($P < 1.3 \times 10^{-15}$), but a significantly higher rate of overlap (9.9%) with the region that was concentrated in the heavy fraction in the log phase ($P < 1.3 \times 10^{-15}$), compared to what would have been expected from random distributions (3.8% and 7.7%, respectively) of the same numbers of nucleotide positions (binomial test with Bonferroni correction).

The same statistical tests were used to evaluate the degree of overlap among the DNA sequences present in high concentrations in the various sucrose fractions in the log and stationary phases. The rate of overlapping nucleotide positions between the genomic regions that were concentrated in the heavy fraction in the stationary phase and those that were concentrated in the histone-rich fraction in the log phase (7.5%), or those that were concentrated in the region enriched in the heavy fraction in the log phase (13.7%), were significantly higher than would have been expected from random distributions of the same numbers of nucleotide positions (3.8% and 7.7% respectively; $P < 1.3 \times 10^{-15}$ in both cases, binomial test with Bonferroni correction).

ACKNOWLEDGMENTS

We thank Hiroki Higashibata (Toyo University) and Shinsuke Fujiwara (Kwansei Gakuin University) for histone cDNAs; Masahiro Tokuhara and Tamotsu Kanai (Kyoto University) for assistance in the manipulation of the *T. kodakaraensis* genome and for helpful suggestions concerning the microarray experiments; Ken Hashida for cloning and sequencing of DNA from the sucrose fractions; Sayaka Iwano for the production of recombinant proteins and antibodies; and the members of the Takeyasu Laboratory for comments on the manuscript. We are also grateful to John N. Reeve (Ohio State University). for insightful suggestions and discussions. This work was supported by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan (T. Itoh, K. Shirahige, T. Imanaka, H. Atomi, S. H. Yoshimura, and K. Takeyasu).

REFERENCES

Ali Azam T, Iwata A, Nishimura A, Ueda S, Ishihama A (1999). Growth phase-dependent variation in protein composition of the *Escherichia coli* nucleoid. J Bacteriol 181, 6361–6370.

Bailey KA, Pereira SL, Widom J, Reeve JN (2000). Archaeal histone selection of nucleosome positioning sequences and the procaryotic

origin of histone-dependent genome evolution. J Mol Biol 303, 25–34.
Bell SD, Botting CH, Wardleworth BN, Jackson SP, White MF (2002). The interaction of Alba, a conserved archaeal chromatin protein, with Sir2 and its regulation by acetylation. Science 296, 148–151.
Bouffartigues E, Buckle M, Badaut C, Travers A, Rimsky S (2007). H-NS cooperative binding to high-affinity sites in a regulatory element results in transcriptional silencing. Nat Struct Mol Biol 14, 441–448.
Brukner I, Sanchez R, Suck D, Pongor S (1995). Sequence-dependent bending propensity of DNA as revealed by DNase I: parameters for trinucleotides. EMBO J 14, 1812–1818.
Dillon SC, Dorman CJ (2010). Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. Nat Rev Microbiol 8, 185–195.
Ehrenhofer-Murray AE (2004). Chromatin dynamics at DNA replication, transcription and repair. Eur J Biochem 271, 2335–2349.
Fujimoto DF, Higginbotham RH, Sterba KM, Maleki SJ, Segall AM, Smeltzer MS, Hurlburt BK (2009). *Staphylococcus aureus* SarA is a regulatory protein responsive to redox and pH that can support bacteriophage lambda integrase-mediated excision/recombination. Mol Microbiol 74, 1445–1458.
Fukue Y, Sumida N, Nishikawa J, Ohyama T (2004). Core promoter elements of eukaryotic genes have a highly distinctive mechanical property. Nucleic Acids Res 32, 5834–5840.
Fukui T, Atomi H, Kanai T, Matsumi R, Fujiwara S, Imanaka T (2005). Complete genome sequence of the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1 and comparison with *Pyrococcus genomes*. Genome Res 15, 352–363.
Fukui T, Eguchi T, Atomi H, Imanaka T (2002). A membrane-bound archaeal Lon protease displays ATP-independent proteolytic activity towards unfolded proteins and ATP-dependent activity for folded proteins. J Bacteriol 184, 3689–3698.
Grainger DC, Hurd D, Goldberg MD, Busby SJ (2006). Association of nucleoid proteins with coding and non-coding segments of the *Escherichia coli* genome. Nucleic Acids Res 34, 4642–4652.
Gribaldo S, Brochier C (2009). Phylogeny of prokaryotes: does it exist and why should we care? Res Microbiol 160, 513–521.
Higashibata H, Fujiwara S, Takagi M, Imanaka T (1999). Analysis of DNA compaction profile and intracellular contents of archaeal histones from *Pyrococcus kodakaraensis* KOD1. Biochem Biophys Res Commun 258, 416–424.
Ishihama Y, Oda Y, Tabata T, Sato T, Nagasu T, Rappsilber J, Mann M (2005). Exponentially modified protein abundance index (emPAI). for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. Mol Cell Proteomics 4, 1265–1272.
Jelinska C, Conroy MJ, Craven CJ, Hounslow AM, Bullough PA, Waltho JP, Taylor GL, White MF (2005). Obligate heterodimerization of the archaeal Alba2 protein with Alba1 provides a mechanism for control of DNA packaging. Structure 13, 963–971.
Kanai T, Akerboom J, Takedomi S, van de Werken HJ, Blombach F, Van Der Oost J, Murakami T, Atomi H, Imanaka T (2007). A global transcriptional regulator in *Thermococcus kodakaraensis* controls the expression levels of both glycolytic and gluconeogenic enzyme-encoding genes. J Biol Chem 282, 33659–33670.
Kim J, Yoshimura SH, Hizume K, Ohniwa RL, Ishihama A, Takeyasu K (2004). Fundamental structural units of the *Escherichia coli* nucleoid revealed by atomic force microscopy. Nucleic Acids Res 32, 1982–1992.
Koide *et al.* (2009). Prevalence of transcription promoters within archaeal operons and coding sequences. Mol Syst Biol 5, 285.
Komori K, Ishino Y (2001). Replication protein A in *Pyrococcus furiosus* is involved in homologous DNA recombination. J Biol Chem 276, 25654–25660.
Koonin EV, Mushegian AR, Galperin MY, Walker DR (1997). Comparison of archaeal and bacterial genomes: computer analysis of protein sequences predicts novel functions and suggests a chimeric origin for the archaea. Mol Microbiol 25, 619–637.
Kouzarides T (2007). Chromatin modifications and their function. Cell 128, 693–705.
Lang B *et al.* (2007). High-affinity DNA binding sites for H-NS provide a molecular basis for selective silencing within proteobacterial genomes. Nucleic Acids Res 35, 6330–6337.
Lee SJ, Surma M, Seitz S, Hausner W, Thomm M, Boos W (2007). Characterization of the TrmB-like protein, PF0124, a TGM-recognizing global transcriptional regulator of the hyperthermophilic archaeon *Pyrococcus furiosus*. Mol Microbiol 65, 305–318.
Lipscomb GL, Keese AM, Cowart DM, Schut GJ, Thomm M, Adams MW, Scott RA (2009). SurR: a transcriptional activator and repressor controlling hydrogen and elemental sulphur metabolism in *Pyrococcus furiosus*. Mol Microbiol 71, 332–349.
Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature 389, 251–260.
Lurz R, Grote M, Dijk J, Reinhardt R, Dobrinski B (1986). Electron microscopic study of DNA complexes with proteins from the Archaebacterium *Sulfolobus acidocaldarius*. EMBO J 5, 3715–3721.
Maeshima K, Eltsov M (2008). Packaging the genome: the structure of mitotic chromosomes. J Biochem 143, 145–153.
Matsunaga F, Forterre P, Ishino Y, Myllykallio H (2001). In vivo interactions of archaeal Cdc6/Orc1 and minichromosome maintenance proteins with the replication origin. Proc Natl Acad Sci USA 98, 11152–11157.
Ohniwa RL, Morikawa K, Kim J, Ohta T, Ishihama A, Wada C, Takeyasu K (2006). Dynamic state of DNA topology is essential for genome condensation in bacteria. EMBO J 25, 5591–5602.
Ohniwa RL, Morikawa K, Takeshita SL, Kim J, Ohta T, Wada C, Takeyasu K (2007). Transcription-coupled nucleoid architecture in bacteria. Genes Cells 12, 1141–1152.
Ohniwa RL, Morikawa K, Wada C, Ohta T, Takeyasu K (2010). Nucleoid architecture and dynamics in bacteria. In: Bacterial DNA, DNA Polymerase and DNA Helicases, ed. SS Bruns, WD Knudsen, New York: Nova Science Publishers, 91–117.
Pereira SL, Grayling RA, Lurz R, Reeve JN (1997). Archaeal nucleosomes. Proc Natl Acad Sci USA 94, 12633–12637.
Pereira SL, Reeve JN (1999). Archaeal nucleosome positioning sequence from *Methanothermus fervidus*. J Mol Biol 289, 675–681.
Perez-Rueda E, Janga SC (2010). Identification and genomic analysis of transcription factors in archaeal genomes exemplifies their functional architecture and evolutionary origin. Mol Biol Evol 27, 1449–1459.
R Development Core Team (2010). R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing: Vienna, Austria.
Robb FT, Place AR (1995). Media for thermophiles. In: Archaea: A Laboratory Manual, ed. FT Robb, AR Place, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 167–168.
Saldanha AJ (2004). Java Treeview—extensible visualization of microarray data. Bioinformatics 20, 3246–3248.
Sandman K, Reeve JN (2005). Archaeal chromatin proteins: different structures but common function? Curr Opin Microbiol 8, 656–661.
Sato T, Fukui T, Atomi H, Imanaka T (2005). Improved and versatile transformation system allowing multiple genetic manipulations of the hyperthermophilic archaeon Thermococcus kodakaraensis. Appl Environ Microbiol 71, 3889–3899.
Schmid AK, Reiss DJ, Pan M, Koide T, Baliga NS (2009). A single transcription factor regulates evolutionarily diverse but functionally linked metabolic pathways in response to nutrient availability. Mol Syst Biol 5, 282.
Schneider KL, Pollard KS, Baertsch R, Pohl A, Lowe TM (2006). The UCSC Archaeal Genome Browser. Nucleic Acids Res 34, D407–D410.
Shioda M, Sugimori K, Shiroya T, Takayanagi S (1989). Nucleosome-like structures associated with chromosomes of the archaebacterium *Halobacterium salinarium*. J Bacteriol 171, 4514–4517.
Smith DR *et al.* (1997). Complete genome sequence of *Methanobacterium thermoautotrophicum* deltaH: functional analysis and comparative genomics. J Bacteriol 179, 7135–7155.
Takayanagi S, Morimura S, Kusaoke H, Yokoyama Y, Kano K, Shioda M (1992). Chromosomal structure of the halophilic archaebacterium *Halobacterium salinarium*. J Bacteriol 174, 7207–7216.
Thanbichler M, Shapiro L (2006). Chromosome organization and segregation in bacteria. J Struct Biol 156, 292–303.
Tomschik M, Karymov MA, Zlatanova J, Leuba SH (2001). The archaeal histone-fold protein HMf organizes DNA into bona fide chromatin fibers. Structure 9, 1201–1211.
Wardleworth BN, Russell RJ, Bell SD, Taylor GL, White MF (2002). Structure of Alba: an archaeal chromatin protein modulated by acetylation. EMBO J 21, 4654–4662.
Woese CR, Kandler O, Wheelis ML (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc Natl Acad Sci USA 87, 4576–4579.
Xue H, Guo R, Wen Y, Liu D, Huang L (2000). An abundant DNA binding protein from the hyperthermophilic archaeon *Sulfolobus shibatae* affects DNA supercoiling in a temperature-dependent fashion. J Bacteriol 182, 3929–3933.
Zimmerman SB (2006). Cooperative transitions of isolated *Escherichia coli* nucleoids: implications for the nucleoid as a cellular phase. J Struct Biol 153, 160–175.